AMENDMENTS TO THE SPECIFICATION

Please amend the specification as shown below:

[0008] The earliest work attempting to characterize and classify the epitopes of particular MHC proteins focused on identifying and screening for anchor residues in epitope peptides and potential epitopes. For example, early methods for prediction focused on characterizing likely epitopes by testing for the presence of the appropriate primary anchors (Falk et al., 1991; Hunt et al., 1992; DiBrino et al., 1993), and secondary anchor residues (Ruppert et al., 1993). An in silico epitope prediction method based on anchor identification was developed by Rammensee et al. (Rammensee et al., 1999). It produced an algorithm for predicting epitopes from protein sequences, and a database (SYFPEITHI~~; http://syfpeithi.bmi-heidelberg.com/~~) of experimentally identified and published motifs, both publicly accessible through a web interface. Elaboration of these techniques lead to the development of EpiMer (De Groot et al., 2001a, which uses a pattern-matching prediction algorithm based on the same principles for identifying peptides that may potentially bind to one or more MHC proteins. Alternatives to these pattern-based methods include neural networks (Gulukota et al., 1997; Milik et al., 1998; Buus, unpublished), statistical methods for parameter estimation (Gulukota et al., 1997), and structure-based methods (Rognan et al., 1994; Altuvia et al., 1997; Rognan et al., 1999; Logean et al., 2001; Schueler-Furman et al., 2001).

[0042] Most of these observations were obtained experimentally, with limited resources and scope. The first approach to large-scale in silico epitope prediction based on anchor identification was taken by Rammensee et al. (Rammensee et al., 1999). It produced an algorithm for predicting epitopes from protein sequences, and a database (SYFPEITHI~~; http://syfpeithi.bmi-heidelberg.com.~~) of experimentally identified and published motifs, both publicly accessible through a web interface. In the algorithm, each amino acid in a candidate peptide is assigned a score that depends on the type of residue, the type of position (anchor, auxiliary), and the frequency of that amino acid at that position in the database of published motifs. The overall peptide score is the sum of the individual amino-acid scores. Ultimately, high-scoring peptides

predicted as immunogenic need to be experimentally validated with binding assays and in vivo and in vitro T-cell assays before becoming good candidates for cancer vaccines.

[0112] In examples shown below, for the A2 allele, five methods are used in the voting heuristic: `AA properties` QP, `BIMAS-like` QP, linear programming, aligrinent (Aln) profile and anchor scoring. For the other alleles that were examined, A1, A3, A24 and B7, four methods are used in the voting heuristic: Parker's method using the most recent matrices maintained at the NIH BioInformatics and Molecular Analysis Section (BIMAS) site ~~(http://bimas.dert.ih.gov/molbio/hla_bind/)~~, and our linear programming, alignment profile and anchors techniques.

TABLE 4 Most favored anchor resides by allele. The preferred values are derived from BIMAS binding matrices ~~(http://bimas.dert.nih.gov/molbio/hla_bind/)~~ developed by the method of Parker et al. (1993).

[0141] For the QP method, a set of 101 HLA-A2 epitopes along with their IC-50 level binding information were extracted from public databases (Parker et al., 1 994a). A set of 694 ninemer epitopes where extracted from the MHCPEP database (~~http://wehih.wehi.edu.au/mhcpep/;~~ Brusic et al., 1998). Of these epitopes, 359 were annotated with their binding strength categories (high, medium or low).

[0144] For the profile-based methods, epitope and ligand sequences previously published for the HLA-A0201 allele were extracted from the SYFPEITHI database (~~http://svlpeithi.bmi-heidelberg.com;~~ Rammensee et al., 1999). Similarly, data for the HLA-A1, A3, A11, A24 and B7 alleles were extracted from the MHCPEP database. After eliminating duplicates and sequences with more or less than nine residues from the 206 HLA-A0201 ligands and epitopes, the remaining 146 distinct ninemers were selected for profile construction. For HLA-A2 anchor scoring, the entire pool of peptides, regardless of length, was analyzed to determine the frequencies of amino-acid pairs at the P2 and C-terminal positions. The procedure for the other alleles was similar.

[0148] In the example below, the algorithms are benchmarked with a reference set of known epitopes and MHC ligand sequences, collected from the literature and from publicly accessible databases. The predictions are compared to those produced with an improved version of the matrix-based approach presented in (Parker et al., 1993), available from the NIH BIMAS site ~~(http://bimas.dcrt.nih.gov/molbio/hla_bind/)~~.